

Abstract

A semantic locator determines whether input sequences form semantically meaningful units. The semantic locator includes a coherence component that calculates a coherence of the terms in the sequence and a variation component that calculates the variation in terms that surround the sequence. A heuristics component may additionally refine results of the coherence component and the variation component. A decision component may make the determination of whether the sequence is a semantic unit based on the results of the coherence component, variation component, and heuristics component.